

**ONLINE SUPPLEMENT
to article in**

AMERICAN SOCIOLOGICAL REVIEW, 2010, VOL. 75

The Temporal Structure of Scientific Consensus Formation

Uri Shwed
Ben Gurion University

Peter S. Bearman
Columbia University

Sensitivity Analysis of Data Extraction

This supplement presents a sensitivity analysis for the data extraction method, based on the longest and most complicated search string—for the case of solar radiation and cancer. Our primary consideration in choosing a search string was maximizing the period covered. We added keywords such as “melanoma” to keywords such as “cancer” to extract more papers from more years and to capture changes in the nomenclature. Here we test the influence of specific search strings on the results, asking whether omission or inclusion of a single term can modify the results.

Ideally, we would like to see a comparison of all possible searches for a specific proposition. This is impossible; just the terms used here can be combined in more than 25,000 different searches, and one can come up with other terms as well. Yet, these results are robust. We show this by comparing several sets constructed by different keywords and examining their differences.

Table S1 shows details of 14 files constructed for this sensitivity analysis. The smallest file (1) is the result of searching for “sun and cancer.” For files 2 through 5, we added terms with an OR connection to the term “sun.” For example, we constructed file 4 by the search “(sun or solar or photo or tanning) and cancer.” We constructed files 6 through 10 by adding terms to the cancer side of the string. File 10 is the one used in the analysis published in the December 2010 issue of *ASR*. We constructed files 11 to 13 by adding more terms to the sun part of the search, and we constructed file 14 by adding “*” to the term cancer. The 14 files cover four periods: file 1 starts in 1992, files 2 through 5 find papers in 1942, 1987, and since 1990. Files 6 through 12 cover 1942 and 1977 through 2005 with the exception of 1980 and 1986, and files 13 and 14 have some papers published in 1986.

We employed the procedures described in the article on each of these files. Then, we calculated Pearson correlations for the modularity scores obtained by every two files in their overlapping years. These correlations are presented in Table S2.

Uri Shwed and Peter S. Bearman
The Temporal Structure of Scientific Consensus Formation

ONLINE SUPPLEMENT
to article in
American Sociological Review, 2010, Vol. 75

Table S1. Files, Search String, and Annual Coverage

File Number, N	Search String	Years in file
1, 1399	Sun & Cancer	1992 – 2005
2, 1877	+ Solar	1942,1987,1990-2005
3, 2194	+Photo	1942,1987,1990-2005
4, 2242	+ Tanning	1942,1987,1990-2005
5, 2527	+ “UV radiation”	1942,1987,1990-2005
6, 3174	+ Melanoma	1942, 1977-9,1981-5, 1987-2005
7, 3611	+ Basal	1942, 1977-9,1981-5, 1987-2005
8, 3639	+ Sarcoma	1942, 1977-9,1981-5, 1987-2005
9, 3822	+carcinoma	1942, 1977-9,1981-5, 1987-2005
10, 4009	+ carci*	1942, 1977-9,1981-5, 1987-2005
11, 4312	+UVA	1942, 1977-9,1981-5, 1987-2005
12, 4832	+UVB	1942, 1977-9,1981-5, 1987-2005
13, 4881	+Sunscreen	1942, 1977-9,1981-2005
14, 4949	+cancer*	1942, 1977-9,1981-2005

ONLINE SUPPLEMENT
to article in

American Sociological Review, 2010, Vol. 75

Table S2: Pearson Correlations of Modularity Scores in Covered Years between 14 Files

1) 1.00 2
2) 0.96 1.00 3
3) 0.97 0.99 1.00 4
4) 0.97 0.99 0.99 1.00 5
5) 0.98 0.99 0.99 0.99 1.00 6
6) 0.95 0.37 0.37 0.38 0.38 1.00 7
7) 0.94 0.39 0.39 0.40 0.40 0.99 1.00 8
8) 0.94 0.35 0.35 0.36 0.35 0.99 0.99 1.00 9
9) 0.94 0.34 0.34 0.34 0.34 0.98 0.98 0.98 1.00 10
10) 0.94 0.40 0.40 0.41 0.40 0.94 0.95 0.95 0.98 1.00 11
11) 0.97 0.38 0.38 0.39 0.38 0.95 0.95 0.95 0.98 0.99 1.00 12
12) 0.97 0.42 0.42 0.43 0.42 0.95 0.95 0.95 0.98 0.99 0.99 1.00 13
13) 0.97 0.41 0.41 0.41 0.41 0.94 0.94 0.94 0.97 0.99 0.99 0.99 1.00 14
14) 0.96 0.43 0.43 0.43 0.43 0.93 0.93 0.93 0.97 0.98 0.98 0.98 0.98 1.00

Table S2 tells a little story. File 1, containing a basic search of “sun and cancer,” has very high correlations with all other files. Files 2 through 5, which include fairly specific keywords regarding sun (connected by OR) but only the word “cancer,” are highly correlated among themselves but show only modest correlation with the other files. These files cover research on a magnitude of things related to the sun and to UV radiation, with some general relation to cancer. For example, one of the most cited papers in file 5 is a paper examining UV radiation as a treatment for prostate cancer.

Files 6 through 14 show extremely high correlations among themselves, as well as with file 1. They cover everything files 1 through 5 had, but they have many more papers that are more specific about the cancer they are researching. It is clear from the correlations that choosing any of these files would yield the same results. But what if a better search for the proposition is one of the more restricted searches, or one that we have failed to try? To examine the actual meaning of a low correlation, Figure S1 plots modularity analysis for file 5 (dashed line), which has the lowest correlation (.4) with file 10 (solid line), which was used in the article.

Figure S1 shows that for the purposes used in this article, even the two least correlated files would yield the same result: modularity peaks (with reference to the observation period covered by both files) in similar years and makes its rapid decline in the same years. Both modularity trends suggest that the literature about the carcinogenicity of the sun’s radiation became consensual in a process that started in 1992 to 1993 and was completed around 1997.

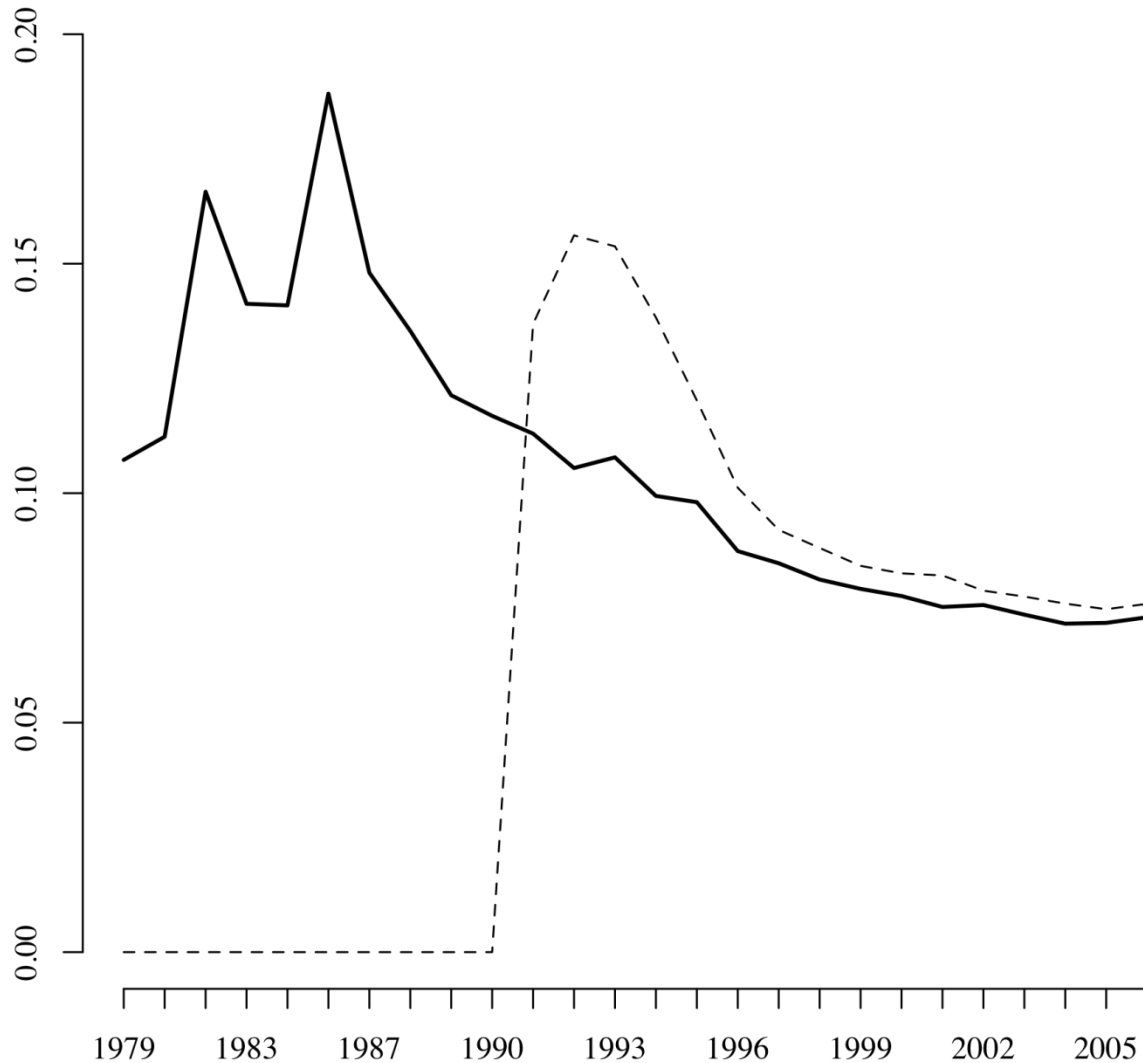


Figure S1. Modularity Analysis of the File Used in the Article with the File Least Correlated

Scaling Modularity to Distinguish between Epistemic Rivalries and Benign Contestation

As science progresses, a literature's size grows. When plotting raw modularity scores over time, we see that modularity usually grows with size, but not always (see Figure S2). Periods in which the literature size grows and modularity declines (e.g., the smoking case in 1958 to 1963 and 1981 to 1985, and the climate change case in 1992 to 1995) show that modularity is not a function of N . Indeed, the way modularity is calculated ignores the number of nodes and controls for the number of edges. So modularity should not be correlated with N . Still, in most years in Figure S2 (and other cases not shown here), modularity and N rise together. Most exceptions are either periods of modularity decline or when modularity is already extremely high.

ONLINE SUPPLEMENT
to article in

American Sociological Review, 2010, Vol. 75

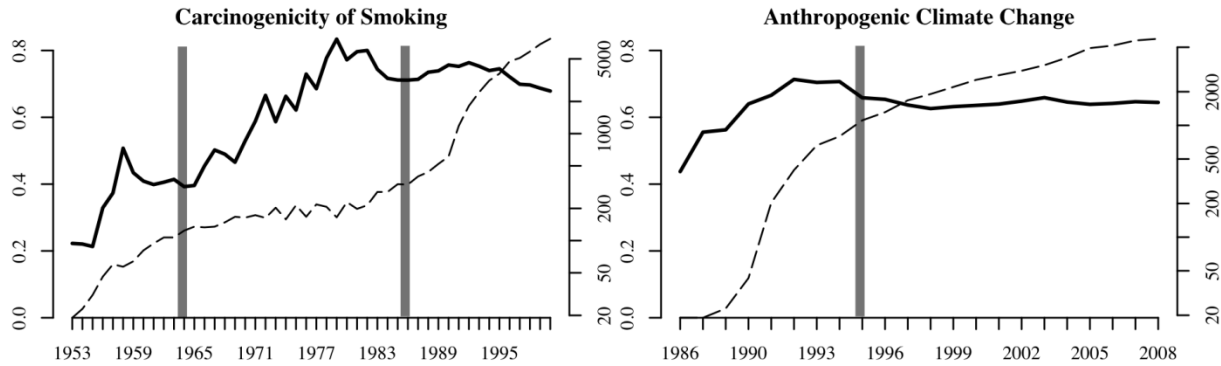


Figure S2: Raw Modularity Scores of Two Cases

Note: Dark lines represent modularity scores, referring to the left-hand-side-axis. Dashed lines represent the number of papers in each observation period, referring to the logarithmic right-hand-side-axis. Grey bars represent crucial consensus reports. Compare to Figure 4 in the article, showing scaled modularity.

While modularity was designed to be robust to network size, some argue that it is affected by it. Indeed, Kreimer and colleagues (2008) argue that modularity increases with size. They examined bacterial metabolic networks and found that modularity increases with network size until a network reaches a size of 109 nodes. In networks of 110 or more nodes there is no correlation between modularity and size. This does not explain the relationship found in our data.

Another possibility is that the specific structure of citation networks creates a different dependency between networks' size and modularity. In citation networks, there is no limit on the number of citations a paper can receive (in-degree), but there is a practical limit on how many papers a given paper can cite. This imbalance may create a correlation between modularity and size. Figure S3 examines whether this can untangle the relationship between size and modularity with 121 simulated networks.

These networks were simulated with a preferential attachment model that resembles citation networks (Barabási and Albert 1999). To accentuate the imbalance between out degree and in degree, we limited the number of citations (ties) a paper (node) can make to three, so that the imbalance is expressed in smaller network sizes. We determine the allocation of ties by a preferential attachment model with a decay function for age, where the probability of an old node (i) to get a new tie is P , determined by its degree k and its age L with a decay function β so that $P[i] \sim (k_i + 1)(L^\beta i)$. This simulates the tendency to cite both widely accepted papers and recent papers. The strength of the tendency to cite recent papers versus old, highly cited papers is governed by the parameter β . We simulated 11 networks of varying sizes for 11 different values of β , between $-.9$ and $-.5$. We plotted the modularity scores of the generated networks against their size in Figure S3.

Figure S3 clearly shows that there is no inherent underlying relationship between modularity and network size. Using a model that describes the evolution of scientific citation networks, modularity is fixed over size. There is a strong correlation, of course, between the decay function in the simulation model and modularity. Decay parameters that are closer to zero produce networks that have a modularity of $.3$ – $.4$, as they penalize old papers less for their age and give rise to huge communities that are merely the product of nodes' degree distribution. When the decay parameter approaches $-.5$, giving only a small chance that an old highly cited paper will be cited more than recent papers, modularity is high as the network produces temporal communities of similar size. However, the overall network size has minimal bearing on the modularity scores. The correlation between modularity and N is $.09$.

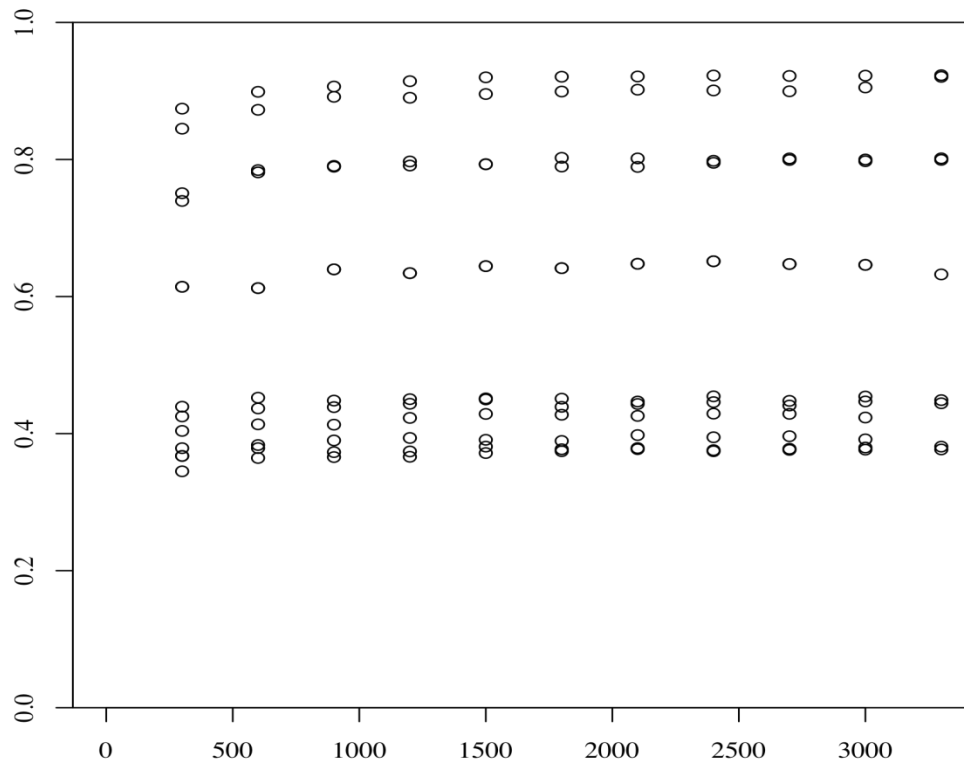


Figure S3. Modularity and Size of 121 Simulated Networks

We can now be confident that networks' size does not confound modularity, but Figure S2 suggests that the two do correlate in the data. Figure S4 describes this relation directly with real data. The figure displays modularity scores and sizes for all of the cases in this article pooled together. By pooling the cases together, we deliberately ignore temporality for the moment. The figure mixes periods of consensus formation with periods of contestation, positioning observations based only on network size and its modularity, to see whether the two are related in our data, even though we know from Figure S3 that it need not be. The top panel of Figure S4 displays raw modularity and raw network size for all of the observation periods in all the cases considered, and the bottom panel displays the same data with logged network size.

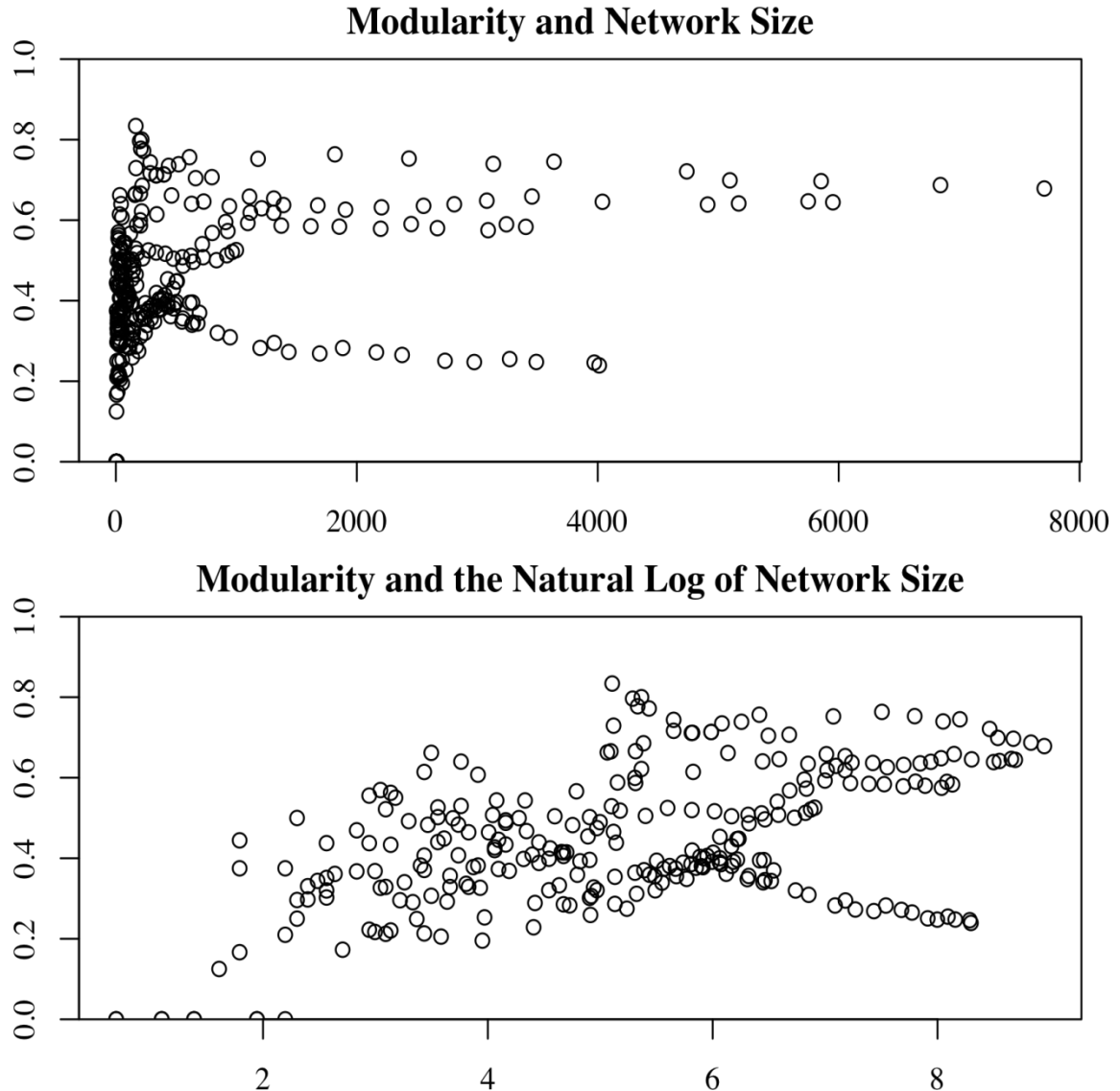


Figure S4. Modularity and Size of 274 Real Citation Networks

Figure S4 shows that modularity and network size are indeed related in the data. There is a general trend of rising modularity with the natural log of network's size. The correlation between modularity and size is .33, and the correlation of modularity with logged network size is .52. In scale-free networks such as scientific citation networks (Barabási and Albert 1999), it makes sense to log the number of nodes. Regardless, both .33 and .52 are strong correlations between size and a measure that is supposed to be robust to size. How is it that simulated networks following a model that resembles citation patterns report a correlation of a mere .09, where the real networks report a correlation as high as .52?

The preferential attachment model does a very good job in explaining network evolution for many different networks, including scientific citation networks. In that sense, it indeed resembles citation

ONLINE SUPPLEMENT
to article in

American Sociological Review, 2010, Vol. 75

networks. But the comparison between Figure S3 (using the model) and Figure S4 (using real data) shows that this model is inappropriate to account for the internal divisions of citations networks. Modularity is not a function of N , mathematically. But it is strongly affected by it, empirically.

Figure S3 shows that size does not mathematically conflate modularity. Figure S2, with its several instances of declining modularity in periods of growth, shows that size does not directly drive modularity in our data. But most of Figures S2 and S4 show that the two are correlated. Figure S2 shows that size is usually correlated to modularity, but not always. Figure S4, stripped of temporal or substantive meaning, shows that the two are correlated when speaking about the abstract object of science.

The reason these two factors are correlated most of the time, but not all of the time, is rooted in what we define as benign contestation. We argue that is beneficial to distinguish between benign contestation—that is, contestation resulting from scientists attempting to establish their own niches in the field—and epistemic rivalries that are large disagreements on core questions. The correlation in Figure S4 describes benign contestation because it is stripped from temporal and case-specific context. Its high correlation of .5 is the result of pooling together networks that describe climate change research in 2005 and tobacco research in 1994, and comparing them with research on the carcinogenicity of coffee from the early 1990s pooled together with tobacco research from the 1960s. These disparate cases and periods do show common logic—expressed in the strong correlation between size and modularity—because in large networks, compared with small networks, there can be many small niches and subdivisions. The correlation between N and modularity—which figure S4 demonstrates is not a mathematical confounding—is simply the expression of benign contestation.

The goal of the published article is to provide a useful measure of the level of epistemic rivalries in scientific literatures, enabling comparative research. Raw modularity does point to consensus formation, but in an obscure manner (see Figure S2). Now we can understand that this is because raw modularity may be composed of benign contestation, which is largely a factor of networks' size, and the epistemic rivalries we care about. Epistemic rivalries are important enough so that when they are resolved, as in the period of consensus formation on the carcinogenicity of tobacco (i.e., the late 1950s and early 1960s), they drive the raw modularity scores down. But when there are no changes in the level of epistemic rivalries, modularity would rise from benign contestation. Understanding this, we control for benign contestation by scaling these scale-free networks. The figures presented in the published article do not report raw modularity, but modularity divided by logged network size, which is a measure of epistemic rivalry.

References

Barabási, Albert-László and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286:509–512.

Kreimer, Anat, Elhanan Borenstein, Uri Gophna, and Eytan Ruppin. 2008. "The Evolution of Modularity in Bacterial Metabolic Networks." 105:6976–81.